Proceedings of the 2021 Winter Simulation Conference
S. Kim, B. Feng, K. Smith, S. Masoud, Z. Zheng, C. Szabo, and M. Loper, eds.

COMPETING INCENTIVES IN SEQUENTIAL SAMPLING RULES

Dashi I. Singham

Operations Research Department
Naval Postgraduate School
1411 Cunningham Road
Monterey, CA 93943, USA

J. George Shanthikumar

Krannert School of Management
Purdue University
403 W. State Street
West Lafayette, IN 47907, USA

ABSTRACT

We describe a framework in which two parties have competing objectives in estimating the mean performance of a stochastic system through sequential sampling. A "regulator" is interested in obtaining a correct measurement with an acceptable probability level, while the "stakeholder" in the project is interested in minimizing the cost associated with the output performance measure and the sampling cost. We demonstrate how the stakeholder can choose an optimal sampling rule to minimize their costs when they have private information which is not shared with the regulator. We further suggest how the regulator can choose a key controllable parameter of the optimal sampling rule in order to asymptotically approach a fair result for both parties.

1 INTRODUCTION

Data collection efforts are important for estimating the performance of a system. Simulation projects are often employed to estimate performance at low cost. However, not all simulation models are inexpensive, with some complex simulation models having extremely long run times or large numbers of design points to test. Additionally, it may be unclear ahead of time how many replications of a simulation are needed to estimate mean performance with a desired level of accuracy because system variance is unknown. Sequential sampling is often employed to minimize the number of replications by simulating until some desired accuracy metric is met.

Because simulation models are not limited by the same costs and constraints of physical experimentation, asymptotic performance of sequential rules is an area of interest. Classical statistics literature has explored the performance of sequential sampling rules, and asymptotic performance results are invoked to justify sequential sampling rules in simulation (Chow and Robbins (1965), Glynn and Whitt (1992)). These sequential rules can be used to ensure, for example, that the accuracy of confidence interval procedures is nominal in the limit.

This paper considers a new perspective on sequential sampling for confidence interval estimation through two parties who use the confidence interval for different purposes. The traditional goal of obtaining accurate confidence intervals is still of interest. However, accuracy may not be the primary goal of a financial stakeholder in a project. In fact, cost minimization may be the most important objective to the stakeholder, who may want to limit the cost spent on delivering a result. We contrast the objectives of a cost minimizing "stakeholder" with an accuracy seeking "regulator" who wants to ensure project performance meets some requirements. For example, an emissions regulator may want to proportionally reward a carbon emitter for reducing their mean emissions level, and they wish to obtain an accurate (with a probability guarantee) estimate of the emissions reduction. The carbon emitter may wish to maximize their expected payoff from the regulator minus costs involved to generate an estimate. The use of sequential rules for regulatory purposes is not new, as in Mukhopadhyay, Bendel, Nikolaidis, and Chattopadhyay (1992) the

978-1-6654-3311-2/21/$31.00 ©2021 IEEE

authors develop a sequential sampling approach to estimate quantiles of acid-neutralizing capability of water from a lake for environmental monitoring purposes.

There are two traditional domains for making sampling decisions. These are the known and unknown variance domains. In the known variance domain, optimal sample sizes and payoffs can be directly derived. In the unknown variance domain, a sequential stopping rule can be used to approximately minimize cost by estimating the variance as samples are collected. This type of cost (or risk) minimization involving a function of both accuracy and sampling cost was explored in Starr (1966a), which proved asymptotic efficiency of risk to the stakeholder. This loosely means that as sampling increases, the risk in the unknown variance case approaches that of the known variance case.

As the variance goes to infinity or the cost of sampling goes to zero, the sample size for sequential rules approaches infinity and the risk to the stakeholder approaches that of the known variance case where the optimal sample size can be derived. However, for a finite variance or a positive replication cost, there is no guarantee that the sequential method will deliver efficient or accurate results, though approximate results can often be shown numerically as in Starr (1966b). These discrepancies from the nominal results appear as the well-known finite-sample bias that exists in sequential procedures. We note that in some cases invariance rules can be derived, whereby an optimal sampling rule can be found that does not depend on the parameters of the underlying distribution (Samuels 1981). We also refer the reader to work on operational statistics (Liyanage and Shanthikumar 2005) which motivates the joint optimization and estimation of key parameters in cases where there is parameter uncertainty. This approach can obtain improved results compared to performing parameter estimation first, then determining the optimal policy.

We introduce a third domain whereby the stakeholder knows their variance but doesn't reveal it to the regulator. The stakeholder must present all data collected for the regulator to calculate a confidence interval but can choose their own sample size, subject to some minimum number of samples set by the regulator. Asymmetric information is a common aspect of two-party contracting models where one side has hidden information, but the other side can design a contract in anticipation of the uncertainty in this private information. See Laffont and Martimort (2002) for a thorough background on principal agent models with asymmetric information. In our setting, the regulator would be the principal and the stakeholder would be the agent holding private information.

This paper attempts to measure the effect of asymmetric information on the performance of sequential rules. If the stakeholder holds this private information about their variance, then they can use a new sequential sampling rule to obtain a reward higher than the optimal case when both parties have equal information. We adopt a framework similar to that in Starr (1966a), and more recently Pasupathy and Yeh (2020), to derive the risk efficiency and expected payoff that can be obtained under asymmetric information. We assume the simulated data is i.i.d. normal as in Starr (1966a) and show a sampling bias may still exist in this case.

Additionally, we show how there may exist a policy available to the regulator in the asymmetric case to approach a limiting fair policy where both parties achieve the known optimal results. Prior papers consider asymptotic risk efficiency by taking the variance or sampling cost to the limit. The regulator cannot know or control the variance or sampling cost, and the stakeholder can benefit from private knowledge of this information. However, the regulator can control the confidence coefficient that will be used to calculate a confidence interval for stakeholder performance. This paper suggests that the regulator can obtain asymptotic efficiency and accuracy as the Type I error α is taken to zero. Thus, it is within the regulator's ability to design a policy that approaches a fair payoff to the stakeholder.

To the best of our knowledge, this paper is the first effort to treat such sampling rules as a two-party problem. We show that the finite-sample and asymmetric information bias exists in an i.i.d. normal data environment, and can potentially be mitigated using a new regulator policy. Section 2 describes the optimal sample size when both the stakeholder and regulator know the underlying variance of the system (the known variance domain), and this case serves as a benchmark for other domains. Section 3 presents the unknown variance domain whereby sequential rules are used when the variance is unknown by either party. Section

4 presents a modified sequential rule in the asymmetric information domain where the stakeholder knows their own variance but keeps it hidden from the regulator. There exists a policy the regulator can use to mitigate the effects of asymmetric information in Section 5. Section 6 concludes and describes potential extensions to more general sampling frameworks.

2 KNOWN VARIANCE DOMAIN

Suppose there is a regulator who wishes to reward a stakeholder for some mean performance level μ. The stakeholder samples their i.i.d. normal performance levels $Y_1, Y_2, \ldots$ sequentially to estimate their mean performance level, with each $Y_i \sim \mathcal{N}(\mu, \sigma^2)$ where σ^2 is the variance of the observations. The regulator is willing to pay a premium proportional to a lower two-sided confidence bound with confidence coefficent $1-\alpha$ for the mean performance level reported, but without loss of generality we set the premium constant to 1. We assume the stakeholder must report all data collected to ensure the confidence bound is calculated correctly, but can choose their own sample size or use any sampling rule. Once the stakeholder reports their data using n samples, the regulator calculates a lower confidence bound of a two-sided confidence interval using a Type I error α, and estimated variance $\hat{\sigma}_n^2$:

$$\overline{Y}_n - t_{\alpha,n-1}\frac{\hat{\sigma}_n}{\sqrt{n}}.$$

The value of $\hat{\sigma}_n^2$ is calculated using the estimate

$$\hat{\sigma}_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \overline{Y}_n)^2.$$

For simplicity, let $t_{\alpha,n-1}$ be short for $t_{1-\alpha/2,n-1} > 0$ which is the quantile of the t-distribution with right-tail probability $\alpha/2$, and similarly denote z_α to be $z_{1-\alpha/2} > 0$ for the standard normal distribution. The stakeholder incurs a linear sampling cost c relative to the payoff from the regulator for each sample collected. Literature such as Ghosh and Mukhopadhyay (1979) considers the relative ratio of the scalar payoff multiplier and the sampling cost, but since we set the payoff multiplier to 1 we assume the cost c is adjusted accordingly. The objective for the stakeholder is to maximize profit (minimize cost):

$$\max_n E\left[\overline{Y}_n - t_{\alpha,n-1}\frac{\hat{\sigma}_n}{\sqrt{n}} - cn\right] \implies \min_n E\left[t_{\alpha,n-1}\frac{\hat{\sigma}_n}{\sqrt{n}} + cn\right]. \tag{1}$$

The simplification of the maximization problem to the minimization holds because $E[\overline{Y}_n]$ is fixed at the true mean μ since it is a martingale. This is because we assume the data is i.i.d. normal and the sampling rules used will depend only on the variance which is independent of the mean. The risk function to the stakeholder from n realized samples is then

$$R_n = t_{\alpha,n-1}\frac{\hat{\sigma}_n}{\sqrt{n}} + cn.$$

For known variance σ^2, by removing integer constraints the optimal choice of n can be derived by translating the right hand side of (1) to the known variance case, taking the derivative, and setting to zero:

$$\begin{aligned}\frac{d}{dn}\left[\frac{z_\alpha\sigma}{\sqrt{n}} + cn\right] &= -\frac{1}{2}z_\alpha\sigma n^{-3/2} + c \implies 0\\ n^* &= \left(\frac{z_\alpha\sigma}{2c}\right)^{2/3}.\end{aligned} \tag{2}$$

Generally, n^* should be at least two to enable a variance estimation. As c approaches zero or σ approaches infinity, the optimal sample size approaches infinity. For the stakeholder, it makes sense that they would want to increase their lower confidence bound towards μ to achieve a higher payoff by reducing the halfwidth via more sampling. On the other hand, as c increases to ∞ or σ decreases to 0, then n^* will trivially be the minimum value of two no matter the observed variance, yielding essentially a fixed-sampling regime. Using the optimal sample size n^*, it can be shown that the optimal value of the risk function is

$$R_{n^*} = 3cn^*.$$

Note also that

$$2cn^* = \frac{z_\alpha \sigma}{\sqrt{n^*}}. \tag{3}$$

Let the optimal payoff made by the regulator to the stakeholder based on the lower bound of the confidence interval be

$$P_{n^*} = \mu - \frac{z_\alpha \sigma}{\sqrt{n^*}}.$$

This payoff is optimal, in that it is the correct/true lower confidence bound that the regulator should pay the stakeholder if all information is known and the optimal sample size is used.

3 UNKNOWN VARIANCE DOMAIN

Given an unknown variance, the stakeholder can use a sequential stopping rule to construct a random stopping time T modeled on the optimal sample size (2). This stopping rule estimates the variance sequentially and determines when enough samples have been collected using:

$$T = \min\left\{ n : n \geq \left(\frac{t_{\alpha,n-1}\hat{\sigma}_n}{2c} \right)^{2/3}, n = m, m+1, \ldots \right\}.$$

Let m be a positive integer that is the minimum sample size that prevents early stopping in the case where the sample variance happens to be estimated close to zero initially. The objective in Starr (1966a) is to derive the minimum starting sample size to ensure asymptotic risk efficiency (to be defined), and given a variance-based payoff and linear cost this minimum starting size is three. In fact, Starr (1966a) considers a more complicated risk function

$$|\overline{Y}_n - \mu|^s - cn^t$$

for some $s,t > 0$ and derives that we need $m > s^2/(s+2t)+1$ for asymptotic risk efficiency. Starr and Woodroofe (1969) then show that the regret (the difference in the risk functions between knowing and not knowing σ) is bounded as a function of σ. Assuming a minimum sample size of three using $s = 2$ and $t = 1$, the stopping rule for T can be rewritten as

$$T = \min\left\{ n : \hat{\sigma}_n \leq \frac{2cn^{3/2}}{t_{\alpha,n-1}}, n = 3, 4, \ldots \right\}, \tag{4}$$

and the risk observed by the stakeholder is

$$R_T = t_{\alpha,T-1}\frac{\hat{\sigma}_T}{\sqrt{T}} + cT.$$

The actual payoff under a random stopping time is

$$P_T = \overline{Y}_T - \frac{t_{\alpha,T-1}\hat{\sigma}_T}{\sqrt{T}},$$

while the expected payoff from the regulator to the stakeholder is

$$E[P_T] = \mu - E\left[\frac{t_{\alpha,T-1}\hat{\sigma}_T}{\sqrt{T}}\right].$$

Define asymptotic risk efficiency when

$$\lim_{c\to 0}\frac{R_T}{R_{n^*}} = 1.$$

This means that as the sampling cost approaches zero, the risk to the stakeholder using the stopping time T approaches that of the true optimal risk. Starr (1966a) establishes asymptotic efficiency results for the i.i.d. normal case (by taking $\sigma \to \infty$ instead of $c \to 0$), while Ghosh and Mukhopadhyay (1979) derives this result for general i.i.d. data under a finite eight moment. Pasupathy and Yeh (2020) extends the work of Ghosh and Mukhopadhyay (1979) to consider dependent non-i.i.d. observations meeting a strong approximation property.

One problem arises in the finite sequential sampling domain when c is not close to 0 or ∞. Then, the risk obtained by the stakeholder will not necessarily be equal to the optimal, and payoffs made by the regulator may be different from the optimal payoff. Note that this bias will exist even under i.i.d. normal data which the stakeholder must report after collecting T observations, so the discrepancies could be even higher for non-i.i.d. or nonnormal data. Starr (1966b) explores what happens in finite-sample contexts for intermediate values of σ/d where d is the required maximum half-width of a confidence interval. In that work, d is analogous to c in (2) in that it appears in the denominator and smaller values increase the sample size towards infinity.

Table 1 displays the numerical performance of stopping rule (4) as $c \to 0$. The simulated data has mean $\mu = 0$ and variance $\sigma^2 = 0.25$, and to calculate the confidence intervals we use $\alpha = 0.1$. We observe that the optimal sample size from the known variance case (n^*) increases. We also observe the average stopping time $\overline{T}$, which is our estimate of $E[T]$, is larger than n^* for smaller sample sizes, but converges to n^* as $c \to 0$. The third row shows the ratio of the average risk at stopping to the optimal risk, and this also converges in the limit. For finite sample sizes, it appears $E[R_T]$ is greater than the optimal risk, which means in the unknown variance case the stakeholder cannot on average achieve the same risk minimization that is possible in the known variance case. Finally, the payoff also converges, and it appears the stakeholder might obtain a higher than optimal payoff for smaller sample sizes on average.

Table 1: Numerical performance of stopping rule T as $c \to 0$ using 1000 replications. The variance is included in parentheses after each term.

c	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
n^*	2.6	11.9	55.3	256.7	1191.4	5529.8
$\overline{T}/n^*$	1.450 (0.067)	1.064 (0.027)	1.015 (0.004)	1.003 (0.001)	1.000 (0.000)	1.001 (0.000)
$\overline{R}_T/R_{n^*}$	1.148 (0.083)	1.013 (0.026)	1.004 (0.004)	1.001 (0.001)	1.000 (0.000)	1.000 (0.000)
$\overline{P}_T/P_{n^*}$	1.020 (0.405)	1.001 (0.040)	1.000 (0.006)	1.002 (0.001)	1.000 (0.000)	1.000 (0.000)

The following proposition derives some bounds on the ratio of the expected risk under an unknown variance sequential rule to that of the known variance optimal risk.

Proposition 1 (unknown variance - risk loss for stakeholder): The ratio of expected achieved risk to optimal risk has the following properties:

1. $\lim_{c\to 0} T/n^* = 1$, $\lim_{c\to 0}\frac{E[R_T]}{R_{n^*}} = 1$;
2. $\frac{E[R_T]}{R_{n^*}} \leq \frac{E[T]}{n^*}$
3. $\lim_{c\to 0}\frac{E[P_T]}{P_{n^*}} = 1$.

Proof: Part 1 is known from Starr (1966a). To understand finite-sample behavior which eventually leads to this convergence, we explore Part 2. To see things from the lower bound perspective, note from Welford (1962) that

$$\hat{\sigma}_T^2 = \frac{T-2}{T-1}\hat{\sigma}_{T-1}^2 + \frac{(Y_T - \overline{Y}_{T-1})^2}{T}.$$

Thus, using the fact that stopping has not happened at time $T-1$ and (4) we have

$$\hat{\sigma}_T^2 \geq \frac{T-2}{T-1}\hat{\sigma}_{T-1}^2 \geq \frac{T-2}{T-1}\frac{4c^2(T-1)^3}{t_{\alpha,T-2}^2},$$

$$\frac{E[R_T]}{R_{n^*}} = \frac{E[t_{\alpha,T-1}\hat{\sigma}_T/\sqrt{T}+cT]}{3cn^*} \geq \frac{E[\lambda 2c(T-1)+cT]}{3cn^*} = \frac{E[(2\lambda+1)T-2\lambda]}{3n^*},$$

where $\lambda = (t_{\alpha,T-1}/t_{\alpha,T-2})\sqrt{(T-2)/T}$. As $c \to 0$, from Part 1 we have $T \to n^*$, and as $T \to \infty$ then $\lambda \to 1$. Thus we have $\lim_{c\to 0} E[R_T]/R_{n^*} \geq 1$. Similarly using (4) we can calculate the upper bound

$$\frac{E[R_T]}{R_{n^*}} = \frac{E[t_{\alpha,T-1}\hat{\sigma}_T/\sqrt{T}+cT]}{3cn^*} \leq \frac{E[2cT+cT]}{3cn^*} = \frac{E[T]}{n^*}.$$

For Part 3, as $c \to 0$, we have $T/n^* \to 1$ as $n^* \to \infty$ (Starr 1966a). Thus, $\hat{\sigma}_T \to \sigma$ and $t_{\alpha,T-1} \to z_\alpha$ and the payoff will be fair asymptotically in that

$$E\left[\frac{t_{\alpha,T-1}\hat{\sigma}_T}{\sqrt{T}}\right] = \frac{z_\alpha \sigma}{\sqrt{n^*}}.$$

□

The stakeholder cannot expect to do better than the optimal fixed sampling rule using n^* by employing a random stopping time using (4). Because they do not know their variance, however, they will still use the sequential stopping rule and collect a suboptimal reward (in expectation). The stakeholder suffers a cost in terms of larger risk by using a sequential rule when they don't know their variance, unless they operate in the limit when $E[T] \to n^*$ using the results from Starr (1966a). Employing the sequential rule (4) will yield a random payoff from the regulator to the stakeholder which could be higher or lower than P_{n^*}, but appears higher in expectation for low n^*.

4 ASYMMETRIC INFORMATION DOMAIN

Next consider the scenario where the stakeholder knows the variance of their output, but this information is hidden from the regulator. One option is for the stakeholder to reveal their variance and use a fixed sampling rule n^* to obtain their optimal risk R_{n^*} and payoff P_{n^*}. Section 3 suggests the stakeholder will do better with this strategy than by using sequential rule (4). However, the stakeholder can potentially do even better than the known optimal by keeping the variance information hidden!

The stakeholder can optimize their risk value using the knowledge of the optimal fixed sample size n^* and obtainable risk R_n^*. To minimize their risk, the stakeholder can use dynamic programming to see if it makes sense to stop early or continue sampling until n^* and accept payoff P_{n^*}, following the general framework of Chow, Robbins, and Siegmund (1971). An optimal sampling rule exists if stopping must stop on or before n^* samples. To use a dynamic programming framework, the boundary condition would be the known payoff R_{n^*} at time n^*. Prior to that, the stakeholder could compare their current risk to their expected risk if they were to continue with an additional sample. The stakeholder may be lucky to observe a lower sample variance early on compared to their known true value, and could result in the best possible opportunity for the stakeholder to minimize their risk.

However, the stakeholder does not even need to use dynamic programming to obtain a smaller than optimal risk result. Here, we consider a simple sequential sampling policy that guarantees a risk less than or equal to R_{n^*} by sampling sequentially until a risk value less than R_{n^*} is achieved, or until the sample size of n^* is reached. Once the sample size n^* is reached, the stakeholder can stop and receive their payoff R_n^*. But if they are lucky in observing an unusually small variance early on, they can save the cost on additional samples and stop with a lower risk. This policy can be written as the following stopping rule $\tilde{T}$:

$$\tilde{T} = \min\left(n^*, \left\{\min\ n : \frac{t_{\alpha,n-1}\hat{\sigma}_n}{\sqrt{n}} + cn \leq 3cn^*, n = 3,4,\ldots\right\}\right) \tag{5}$$

which can be rewritten as

$$\tilde{T} = \min\left(n^*, \left\{\min\ n : \hat{\sigma}_n \leq \frac{(3n^*-n)c\sqrt{n}}{t_{\alpha,n-1}}, n = 3,4,\ldots\right\}\right). \tag{6}$$

We make the following claims about the performance of the asymmetric information sampling rule $\tilde{T}$.

Theorem 1 (asymmetric information): Assume $0 < \sigma^2 < \infty$. Under sampling rule (6) with stopping time $\tilde{T}$, the stakeholder achieves a lower risk value than the known optimal for finite stopping times and equal risk and payoff values asymptotically, i.e.,

1. For $c > 0$, $E[\tilde{T}] \leq n^*$, $R_{\tilde{T}} \leq R_{n^*}$, and $E[R_{\tilde{T}}] \leq R_{n^*}$;
2. $\lim_{c\to 0}\tilde{T}/n^* = 1$ and $\lim_{c\to 0}E[\tilde{T}]/n^* = 1$;
3. $\lim_{c\to 0}E[P_{\tilde{T}}]/P^* = 1$.
4. $\lim_{c\to 0}E[R_{\tilde{T}}]/R_{n^*} = 1$;

Proof: Part (1) follows immediately by construction of the rule. The values of $\tilde{T}$ must be less than or equal to n^*, and the stakeholder always has the option to wait until time n^* to collect R_{n^*}, so will achieve at most that level of risk. Part (2): We follow a similar approach to Theorem 3 of Pasupathy and Yeh (2020). Without a limit on the stopping time at n^*, it is not guaranteed that the condition

$$\hat{\sigma}_n \leq \frac{(3n^*-n)c\sqrt{n}}{t_{\alpha,n-1}}, n = 3,4,\ldots$$

will ever be true because $\frac{(3n^*-n)c\sqrt{n}}{t_{\alpha,n-1}}$ can become negative quickly in which case the stopping condition on $\hat{\sigma}_n$ will never be reached, and hence stopping will occur at n^*. Note in (5), as $c \to 0$, cn goes to zero much faster than $3cn^*$ which is $O(c^{1/3})$. Then in the limit $\tilde{T}$ is approximately the first time $\frac{t_{\alpha,n-1}\hat{\sigma}_n}{\sqrt{n}} \leq 3cn^*$, which is when $n \geq (t_{\alpha,n-1}\hat{\sigma}_n/3cn^*)^2$, or $\tilde{T} = n^*$ if the stopping condition is not reached by n^*. If stopping before n^* happens, then we have

$$\frac{t_{\alpha,\tilde{T}-1}\hat{\sigma}_{\tilde{T}}}{3cn^*} \lesssim \tilde{T} \lesssim \frac{t_{\alpha,\tilde{T}-2}\hat{\sigma}_{\tilde{T}-1}}{3cn^*}.$$

Dividing both sides by n^* and taking the limit yields terms cancelling so $1 \lesssim \lim_{c\to 0}\tilde{T}/n^* \lesssim 1$. Parts (3) and (4) follow from Part (2). If $\tilde{T} \geq n^*$ then $P_{\tilde{T}}$ and $R_{\tilde{T}}$ will equal P_{n^*} and R_{n^*} respectively. If $\tilde{T} \leq n^*$, we know from Part (2) that $\tilde{T}/n^*$ approaches 1 so the payoff and risk functions will be optimal in the limit. □

Thus, under asymmetric information, stopping rule (6) has an asymptotically fair result, in that both sides achieve the known optimal case. Additionally, in contrast to the unknown variance result from Proposition 1 which appears biased against the stakeholder, Part (1) of Theorem 1 shows that it is possible for the stakeholder to construct a policy under asymmetric information which is biased in their favor.

5 OPTIMAL REGULATOR POLICY UNDER ASYMMETRIC INFORMATION

In order to prevent the stakeholder benefiting from asymmetric information, the regulator needs to account for the potential existence of bias in a finite-sampling environment. The regulator cannot control or know c or σ, so while limiting results are fine when the interests of the stakeholder and regulator are aligned, in practice the regulator cannot control these variables. One option is to increase the payoff scaling factor so that this effectively diminishes c in the eyes of the stakeholder, but pushing this payoff to infinity yields high costs for the regulator. However, α can be controlled by the regulator and set as a specification for the lower confidence bound. This setting becomes a game, whereby the regulator must choose some α ahead of time knowing the stakeholder will try to minimize risk at the expense of potentially reporting an inaccurate confidence bound.

The problem of risk bias in this setting is closely related to that of the well-known confidence interval coverage problem in absolute and relative-precision stopping rules. These rules deliver confidence intervals with nominal coverage asymptotically as the desired half-width approaches zero. However, coverage is usually less than desired for finite stopping times in sequential rules. Singham and Schruben (2012) show that rather than decreasing the half-width to zero and approaching nominal coverage from below, it is optimal to inflate the coverage coefficient $1-\alpha$ (decreasing α) to actually obtain nominal coverage while minimizing sampling. We suggest the following theorem for the asymptotic efficiency with respect to α when stopping rule T is used (both parties have unknown variance).

Theorem 2 (asymptotic regulator policy - unknown variance). Assume $0<\sigma^2<\infty$ and $c>0$, then,

1. $\lim_{\alpha\to 0} T=\infty$ a.s. and $\lim_{\alpha\to 0} E[T]=\infty$;
2. $\lim_{\alpha\to 0} T/n^*=1$ a.s. and $\lim_{\alpha\to 0} E[T]/n^*=1$;

Proof: Part (1): Referring back to (4), as $\alpha\to 0$, $t_{\alpha,n-1}\to\infty$ for all values of n. The result follows immediately, for example, using Lemma 1 of Chow and Robbins (1965). Part (2): From parts (4) and (6) of Theorem 1 of Chow and Robbins (1965). □

Because it is difficult to establish conclusively the performance of stopping rule $\tilde{T}$ with respect to α, we show numerical performance that can be obtained when $\alpha\to 0$ and compare it to that obtained by decreasing $c\to 0$. We simulate data observations Y_i with mean $\mu=1$ and variance $\sigma^2=0.25$. For 1,000 replications, we simulate until the stopping rule is met, and collect $\tilde{T}, R_{\tilde{T}}$, and $P_{\tilde{T}}$. First we show the effect of decreasing c using a fixed $\alpha=0.1$ in Table 2, where we report the estimated risk and payoff compared to the optimal values.

Table 2: Numerical performance of stopping rule $\tilde{T}$ as $c\to 0$ using 1,000 replications. The variance is in parentheses after each result.

c	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
n^*	2.6	11.9	55.3	256.7	1191.4	5529.8
$\overline{\tilde{T}}/n^*$	1 (0)	0.700 (0.089)	0.708 (0.095)	0.831 (0.032)	0.887 (0.017)	0.932 (0.006)
$\overline{R}_{\tilde{T}}/R_{n^*}$	1 (0)	0.923 (0.020)	0.980 (0.006)	0.998 (0.001)	0.998 (0.001)	1.000 (0.000)
$\overline{P}_{\tilde{T}}/P_{n^*}$	1 (0)	0.997 (0.057)	0.985 (0.012)	0.996 (0.001)	1.000 (0.001)	1.000 (0.000)

For the large values of c in the first column of Table 2, the optimal known stopping value effectively happens at the minimum number of samples. As this is a fixed sampling situation, the risk and payoff are the same as the optimal so the ratios are exactly 1. As c approaches 0, the sample size increases and the stopping time approaches n^*. Additionally, we see the risk and payoff approach their optimal values as expected. We next explore the effect of decreasing α. Fix the sampling cost at $c=0.001$. Table 3 displays the results as α is decreased towards zero. The optimal sample size increases, but not nearly as quickly as when c decreases to 0.

Table 3: Numerical performance of stopping rule $\tilde{T}$ as $\alpha \to 0$ using 1,000 replications. The variance is in parentheses after each result.

α	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
n^*	55.3	75.6	87.8	98.2	106.8	114.4
$\overline{\tilde{T}}/n^*$	0.699 (0.091)	0.806 (0.059)	0.855 (0.040)	0.881 (0.033)	0.909 (0.025)	0.928 (0.018)
$\overline{R}_{\tilde{T}}/R_{n^*}$	0.978 (0.007)	0.993 (0.001)	0.998 (0.000)	0.998 (0.000)	0.999 (0.000)	0.999 (0.000)
$\overline{P}_{\tilde{T}}/P_{n^*}$	0.985 (0.014)	0.990 (0.007)	0.988 (0.004)	0.985 (0.003)	0.990 (0.002)	0.989 (0.002)

We still observe convergence in the value of $\tilde{T}$ and the risk and payoff functions. The results suggest a similar conclusion to Singham and Schruben (2012), in that it is more efficient to decrease α than c to obtain asymptotically valid results while increasing the total number of samples at a much slower rate. In addition to α being a parameter that is controllable by the regulator, it may lead to results that are closer to the optimal for both parties without needing to greatly increase the number of samples.

6 CONCLUSION

This paper introduces the notion of asymmetric information into the standard sequential sampling framework. One party, the regulator, is interested primarily in the accuracy of the reported results through delivery of a fair payoff based on the lower confidence bound. A second party, the stakeholder, is interested in minimizing their risk based on the received payoff and sampling cost, and so has a different objective than the regulator. When both parties have the same information, either the optimal result can be delivered in the known variance case, or an asymptotically fair policy can be constructed as the cost parameter goes to zero in the unknown variance case. However, under non-asymptotic conditions, the regulator cannot control the cost parameter and so the payoff may be unfair relative to the optimal, while the stakeholder likely achieves risk levels higher than optimal. When the stakeholder knows their variance and keeps it hidden from the regulator, they can construct an improved stopping rule that will always deliver lower risk levels than optimal.

We identify a parameter the regulator can control and push to the limit to obtain a fair policy in both the unknown variance and asymmetric information domain. This parameter is the Type I error associated with a confidence coefficient. This type of analysis has potential to deliver limiting efficiency in many other sequential sampling contexts, for example, confidence interval coverage. Further, we hypothesize that the limiting results can be reached with fewer samples using the Type I error rather than the traditional approach of decreasing the sampling cost. Future work will compare the rates of convergence for these parameters to assess the sampling benefits. Additionally, a backwards-induction method can be used to potentially find an optimal sampling rule for the stakeholder in the asymmetric information domain that outperforms the one introduced here.

ACKNOWLEDGMENTS

We are extremely grateful to Raghu Pasupathy for helpful feedback and comments on an early version of this paper.

REFERENCES

Chow, Y., and H. Robbins. 1965. “On the Asymptotic Theory of Fixed-Width Sequential Confidence Intervals for the Mean”. *Annals of Mathematical Statistics* 36(2):457–462.

Chow, Y., H. Robbins, and D. Siegmund. 1971. *Great Expectations: The Theory of Optimal Stopping*. New York: Houghton Mifflin.

Ghosh, M., and N. Mukhopadhyay. 1979. “Sequential Point Estimation of the Mean when the Distribution is Unspecified”. *Communications in Statistics-Theory and Methods* 8(7):637–652.

Glynn, P., and W. Whitt. 1992. "The Asymptotic Validity of Sequential Stopping Rules for Stochastic Simulations". *Annals of Applied Probability* 2(1):180–198.

Laffont, J.-J., and D. Martimort. 2002. *The Theory of Incentives: The Principal-Agent Model*. 2nd ed. Princeton, NJ: Princeton University Press.

Liyanage, L. H., and J. G. Shanthikumar. 2005. "A Practical Inventory Control Policy Using Operational Statistics". *Operations Research Letters* 33(4):341–348.

Mukhopadhyay, N., R. B. Bendel, N. P. Nikolaidis, and S. Chattopadhyay. 1992. "Efficient Sequential Sampling Strategies for Environmental Monitoring". *Water Resources Research* 28(9):2245–2256.

Pasupathy, R., and Y. Yeh. 2020. "Risk-Efficient Sequential Simulation Estimators". In *Proceedings of the 2020 Winter Simulation Conference*, edited by K.-H. Bae, B. Feng, S. Kim, S. Lazarova-Molnar, Z. Zheng, T. Roeder, and R. Theising, 2879–2886. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Samuels, S. M. 1981. "Minimax Stopping Rules when the Underlying Distribution is Uniform". *Journal of the American Statistical Association* 76(373):188–197.

Singham, D., and L. Schruben. 2012. "Finite-Sample Performance of Absolute-Precision Stopping Rules". *INFORMS Journal on Computing* 24(4):624–635.

Starr, N. 1966a. "On the Asymptotic Efficiency of a Sequential Procedure for Estimating the Mean". *The Annals of Mathematical Statistics* 37(5):1173–1185.

Starr, N. 1966b. "The Performance of a Sequential Procedure for the Fixed-Width Interval Estimation of the Mean". *The Annals of Mathematical Statistics* 37(1):36–50.

Starr, N., and M. B. Woodroofe. 1969. "Remarks on Sequential Point Estimation". *Proceedings of the National Academy of Sciences* 63(2):285–288.

Welford, B. 1962. "Note on a Method for Calculating Corrected Sums of Squares and Products". *Technometrics* 4(3):419–420.

AUTHOR BIOGRAPHIES

DASHI I. SINGHAM is a Research Associate Professor in the Operations Research Department at the Naval Postgraduate School where she researches, teaches, and advises student theses. Her main research interests lie in simulation analysis, methodology and applied statistics, and she has worked in the areas of sequential stopping rules for simulation output analysis, simulation optimization, and standardized time series. She obtained her Ph.D. in Industrial Engineering & Operations Research from the University of California at Berkeley and her B.S.E. in Operations Research and Financial Engineering from Princeton University. Her email address is dsingham@nps.edu, and her website is https://faculty.nps.edu/dsingham/.

J. GEORGE SHANTHIKUMAR is is Distinguished Professor of Management at Purdue University. His research interests are in integrated inter-disciplinary decision making, model uncertainty and learning, production systems modeling and analysis. His email address is shanthikumar@purdue.edu. His website is https://web.ics.purdue.edu/∼jshanthi.